

Les orthographes d'Arsène Vincent, chroniqueur et copiste (Perche d'Eure-et-Loir aux environs de 1870) Informatisation d'un manuscrit et variation graphique

Introduction

Avec le volumineux manuscrit portant sur la page de titre *Recherches historiques sur le canton de Thiron-Gardais, par Ar. Vincent, sculpteur*¹, nous pénétrons dans une nébuleuse orthographique dont les formes et les systématiques d'écriture n'ont pas fini de nous surprendre, tant leur espace de variation est étendu, en raison de l'insécurité graphique de l'auteur et de la diversité des sources reproduites. Ce document s'avère être du plus grand intérêt orthographique par la qualité et la quantité de données qu'il apporte. À travers lui, en particulier, il semble possible de pouvoir éclaircir un peu cet univers de l'orthographe peu assurée d'un citoyen ordinaire, avant l'avènement de l'école obligatoire.

Arsène Vincent (1831-1881) est un ancien charron devenu historien local après l'accident qui le paralysa². Son ouvrage manuscrit, resté inachevé à la lettre P, se présente sous forme de dictionnaire avec différentes rubriques classées par ordre alphabétique (788 entrées). Ce *Dictionnaire*, ainsi nommé dorénavant, a dû être écrit autour des années 1870, ce qui est attesté par des dates et un formulaire de « Dénombrement de 1876 », inséré à la fin du volume et au dos duquel figure un article. Avant le *Dictionnaire*, aux alentours de 1864, il a composé un *Glossaire* réunissant quelques 432 termes patois (voir, ci-dessous l'*Annexe 1 : Données sur le Dictionnaire et le Glossaire*).

La variation et les contrastes graphiques entre le *Dictionnaire* et le *Glossaire* sont illustrés par des articles communs aux deux écrits et comparés aux graphies des entrées du *Trésor du parler percheron* (abréviation *TPP*). D'une manière générale, le *Dictionnaire* est plus soigné que le *Glossaire*, non seulement dans la composition manuscrite, mais aussi dans l'écrit qui est plus surveillé et présente de nombreux hypercorrectismes. Les différentes graphies des entrées, comme les gloses d'Arsène Vincent donnent rapidement un aperçu des problématiques de la variation graphique où écrit dialectal et écrit français peu assuré sont réunis dans un même espace graphique (voir, ci-dessous l'*Annexe 2 : Bref aperçu : contrastes articles du Dictionnaire / articles du Glossaire*).

L'exposé qui suit présente deux parties. La première traite de la méthode d'informatisation du corpus manuscrit. La seconde partie aborde la variation graphique avec les premiers résultats d'une recherche qui débute seulement et qui est appelée à de nombreux développements.

¹ Chef-lieu de canton de l'Eure-et-Loir est situé entre Nogent-le-Rotrou et Chartres (15 kms de Nogent, 40 kms de Chartres et 130 kms de Paris).

² Pour les renseignements sur l'auteur et son oeuvre cf. Marie-Rose Simoni-Aurembou, *L'alphabet du quotidien*, 1982, pp. 3-7.

1. Simulation informatisée du manuscrit

Le concept de simulation informatisée du manuscrit est construit autour de l'idée de reproduire un document informatique analogue au document original de référence. Cette opération consiste non seulement à saisir intégralement le texte, mais invite également à tenir compte de l'ensemble des éléments matériels qui participent à son élaboration et à sa sémiotique graphique globale. Ce point de vue concerne à la fois la micro et la macrostructure de l'ouvrage, c'est-à-dire la mise en écrit du texte, sa mise en page et sa structuration dynamique tout au long du volume.

L'objectif d'une telle simulation est de mettre en évidence et de reconnaître chacune des composantes de l'ouvrage comme élément d'information d'un système fonctionnant en réseau. Ainsi, l'utilisation de la couleur permet de marquer la diversité des sources entre parties autographes et parties reproduites, tout comme elle permet aussi d'isoler le système de renvois et les autres éléments organisationnels du texte, participant à l'économie générale de l'ouvrage.

Ce premier système de marquage par la couleur, qui relève globalement de la composition et de la structure du dictionnaire, est complété par un second système de marquage destiné à enregistrer tous les écarts graphiques et toutes les modifications du manuscrit. En effet, l'analyse informatisée des écarts et des modifications, en apportant une meilleure connaissance des données, devrait permettre de progresser dans l'étude du travail d'écriture de l'auteur et de la genèse de son système graphique.

Cet usage de la couleur, qui est visualisé à l'écran et peut être imprimé sur support papier, facilite les travaux de l'analyste. C'est un moyen simple à mettre en place et à utiliser, il peut être assimilé à un système spécifique de balisage du texte. Mais, c'est avant tout par la constitution d'un ensemble de ressources informatiques que l'on saisit l'intérêt d'une telle démarche pour une analyse de la variation graphique qui traite aussi bien les aspects qualitatifs que quantitatifs.

La suite de l'exposé présente la mise en place de ce système et les premiers résultats qu'il a permis d'obtenir.

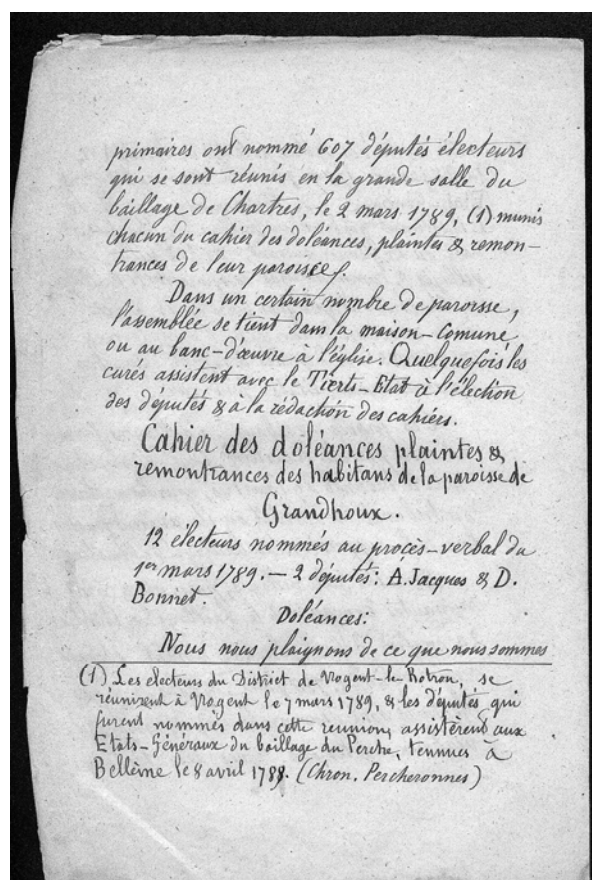
1.1. Marquage des sources et du système de renvois

Dans le corpus retenu (lettre A), en dehors des parties autographes et du système de renvois, l'auteur y reproduit quatre documents de sources différentes et de chronologies différentes³. À chacune de ces sources, comme à chacune de ces chronologies, correspond un système graphique particulier et, l'usage de la

³ 1 : *Cahier de doléances plaintes et remontrances des habitants de la paroisse de Grandhoux, de la paroisse de La Gaudaine et de Marolles* (1789, pp 8-18) ; 2 : *Revue de l'agriculture percheronne* (1849, pp 19-32) ; 3 : Jean de Troyes, *Chronique de Louis XI*, sous l'article *Angelus* (1472, pp. 50-51) ; 4 : Discours du Curé Lépine, sous l'article *Arbre-de-liberté* (1848, pp 74-78). Chacun de ces documents présente ses propres caractéristiques qui, en dehors de quelques graphies, semblent relativement en adéquation avec les usages graphiques des différentes époques.

couleur, associé à chacune de ces différences, permet le maintien de toutes ces distinctions.

Ce marquage par la couleur est présenté par l'exemple, ci-dessous⁴, où le texte de la page manuscrite originale est reproduit dans la colonne de gauche et sa simulation informatisée dans la colonne de droite. Dans l'illustration en noir et blanc, la distinction faite par les couleurs est compensée par différents types de barres verticales à la marge gauche du texte imprimé. Ainsi : la barre mince simple indique une partie autographe ; la barre mince double signale la partie reproduite du *Cahier des doléances* ; la barre simple épaisse désigne le système de renvois, ici une note de bas de page.



primaires ont nommé 607 députés électeurs
qui se sont réunis en la grande salle du
baillage de Chartres, le 2 mars 1789, (1) munis
chacun du cahier des doléances, plaintes & remon-
trances[CP] de leurs **paroiss**es.

Dans un certain nombre de **paroisse**,
l'assemblée se tient dans la **maison-Comune**
ou au banc-d'oeuvre à l'église. Quelquefois les
curés assistent avec le **Tierce-Etat** à l'élection
des députés & à la rédaction des cahiers.

Cahier des doléances plaintes &
remontrances des **habitans** de la paroisse de
Grandhoux.

12 électeurs nommés au procès-verbal du
1^{er} mars 1789. - 2 députés : A. Jacques & D.
Bonnet.

Doléances :
Nous nous plaignons de ce que nous sommes

(1) Les **electeurs** du District de Nogent-le-Rotrou, se
réunirent à Nogent le 7 mars 1789, & les députés qui
furent nommés dans cette **reunion**, assistèrent aux
Etats-Généraux du baillage du Perche, **tennues** à
Bellême le 8 avril 1789[-8]. (Chron. Percheronnes)

Enfin, les mots en caractères gras indiquent des formes présentant des écarts graphiques ou des modifications du manuscrit. Il va de soi que les écarts graphiques sont déduits à partir des synchronies respectives relatives aux divers documents.

1.2. Index global avec distinction des parties

À partir de l'utilisation de la couleur telle qu'elle vient d'être décrite, il est possible, par filtrages successifs - à la manière d'un photographe procédant à la séparation des couleurs -, de générer un index global où apparaissent les formes

⁴ L'exemple présenté correspond à la page 8 du manuscrit.

de chacune des parties distinctes. Les tableaux, ci-dessous, donnent un aperçu des résultats obtenus grâce à cette démarche :

<i>variantes entre sources</i>			<i>variantes dans même source</i>		
<i>forme</i>	<i>source</i>	<i>occ.</i>	<i>forme</i>	<i>source</i>	<i>occ.</i>
pourraient	Rev. Agric.	1	Ansuite	Glos.	1
pourrait	Rev. Agric.	1	ensuite	Glos.	2
pourait	Glos.	1	ensuite	Glos.	1
pourrait	Glos.	5	l'~ors~qu'	Glos.	1
pouoient	Cah. Dol.	1	l'~orsqu'	Glos.	1
pouroit	Cah. Dol.	2	l'~orsque	Glos.	1
apparoître	Cah. Dol.	1	lors~qu'	Glos.	1
renoître	Cah. Dol.	1	lors~que	Glos.	2
hiver	Glos.	1	l'~orsqu#<'>un	Glos.	1
hyver	Cah. Dol.	1	lorsqu#il	Glos.	1
roi	Glos.	3	lorsqu'	Glos.	4
Roy	Cah. Dol.	3	lorsque	Glos.	4

Les formes **en caractères gras** signalent des écarts graphiques en rapport avec les usages de l'époque

Abréviations : Cah. Dol. = Cahiers de Doléances (mars 1789) ; Glos. = Gloses des articles ; occ = nombre d'occurrences de la forme ; Rev. Agric. = Revue de l'Agriculture percheronne (1849)

- À gauche, dans le tableau *variantes entre sources*, les graphies provenant du *Cahier de doléances* dénotent des usages anciens, voire archaïsants : le phonème [E] est encore noté par *oi*, le graphème *oi* n'a été remplacé par *ai* que dans la cinquième édition du *Dictionnaire de l'Académie française* (1835) ; de même, la forme *hiver*, bien qu'enregistrée avec *i* dès *Académie* 1762, est encore notée **hyver** (avec y) ; idem pour la forme **Roy**, avec maintien du y (alors que la forme moderne *roi* est enregistrée dès *Académie* 1740) et où l'usage de la majuscule, à toutes les occurrences, exprime le respect envers le monarque. Toutes les autres occurrences du tableau, dont celles de l'auteur (Glos.), dénotent des formes modernes encore en usage aujourd'hui, en dehors de l'écart graphique de **pourait** où la consonne double est écrite par *r* simple.
- À droite, le tableau *variantes dans même source* (tous les exemples proviennent de parties autographes), révèle des écarts graphiques de la forme *ensuite* où la voyelle nasale est parfois notée par **an** à la place de *en*, c'est une hésitation récurrente chez l'auteur et on constate souvent ce type d'écart graphique pour beaucoup d'autres mots ; sur le même mot, le doublement de la consonne de la syllabe finale (*-suite*, 2 occurrences, pour *-suite*, 1 occurrence) traduit une tendance fréquente à l'hypercorrectisme. Enfin, la forme *lorsque*, avec pas moins de 7 formes graphiques différentes, présente de multiples écarts de segmentation (le tilde ~ indique une déglutination, le dièse # indique une agglutination). Et, malgré l'impression laissée par ces nombreuses variantes, l'usage graphique, entre les formes présentant un écart (8 occurrences) et celles n'en présentant pas (8 occurrences), reste équilibré. Les explications de ces

écarts restent à découvrir, une analyse approfondie et systématique de tous les contextes permettrait probablement de progresser dans l'interprétation.

1.3. Marquage des modifications et des faits matériels du manuscrit

L'excellente conservation du manuscrit, liée à la qualité du papier et de l'encre, ouvre une intéressante perspective de recherche, non seulement sur les hésitations graphiques de l'auteur, mais également sur les contraintes matérielles dues à la structure sous forme de dictionnaire de l'ouvrage. Du point de vue du système graphique, la plupart des modifications ont laissé des traces visibles et déchiffrables qui permettent, par la confrontation entre l'état original et l'état final, de suivre les tâtonnements de l'écriture. Et d'autre part, sous l'aspect de la structure de l'ouvrage, les notes de bas de page, les annotations en marge et les espaces blancs réservés à la notation de renvois internes ou à des compléments d'informations, sont le témoignage d'une conception pratique et méthodique du volume. Le tableau, ci-dessous, présente le codage de chacun de ces éléments matériels :

	<i>système de marquage</i>	<i>exemple ou explication</i>
1	addition : < >	ex<c>ellence
2	substitution : { }	allait{-er}
3	suppression : —	même (caractères barrés)
4	segmentation : # et ~°	quelqu#in<s> (agglutination)
5		quel~°qu#<'>un (déglutination)
6	caractères : \$	atteindre{\$\$\$} (illisibilité)
7	caractère : £	coeur£ (symbole calligraphique)
8	souligné continu	signale un effet graphique de mise en relief
9	souligné discontinu	signale les mots et expressions situés dans la marge
10	[BDR]	indique un blanc de réserve (complément et/ou renvoi)
11	omission ou restitution : []	signale un mot manquant ou une ponctuation restituée

• Les lignes numérotées de 1 à 6 concernent plutôt le marquage des modifications graphiques du manuscrit :

- 1 - les soufflets (symboles < >), encadrent un élément ajouté (addition). Ainsi (cf. l'exemple de la ligne 1), à l'état initial **excellence** est ajouté un **c** entre **x** et **e** dont l'état final **ex<c>ellence** témoigne qu'il y a eu hésitation avant l'aboutissement à la graphie sans écart ;
- 2 - les accolades (symboles { }), délimitent ce qui a été substitué à l'état original (substitution). Ainsi (cf. l'exemple de la ligne 2), à l'état initial **aller** (qui note la 3^e personne de l'imparfait) le **-er** final, dont il subsiste la trace sur le manuscrit, est remplacé par **-ait** suffixe verbal de la 3^e personne de l'imparfait ;

- 3 - l'attribut de caractère **style barré** est utilisé pour les mots et expressions supprimés (cf. l'exemple de la ligne 3) où l'occurrence **même** indique la suppression de ce mot ;
- 4 - le dièse (symbole #) représente l'agglutination d'éléments normalement séparés dans l'usage graphique usuel. Ainsi (cf. l'exemple de la ligne 4), la forme **quelqu#in<s>** signifie que le mot a d'abord été écrit **quelquin**, état original, avec agglutination des deux éléments et notation de la voyelle nasale **un** par **in**. Et, dans un second temps, l'auteur a ajouté un -s final notant le pluriel, probablement par hypercorrectisme. La forme **quelquins** de l'état final, grâce à ce système de transcription est noté par la forme, **quelqu#in<s>**, qui permet de reconstituer pas à pas toutes les étapes de l'écriture dont l'accumulation des écarts manifeste bien l'insécurité orthographique et les hésitations de l'auteur ;
- 5 - le tilde et le petit o en exposant (symbole ~°) désignent la déglutination d'une unité en deux éléments graphiques. Ainsi (cf. l'exemple de la ligne 5), les composants de l'expression **quel~°qu#<'>un** correspondent d'abord à la séquence **quel quun** de l'état original qui marque à la fois une séparation graphique des deux premiers éléments et une agglutination graphique des deux derniers éléments. Ensuite, l'auteur a ajouté une **apostrophe** pour indiquer que le second terme est constitué de deux éléments graphiques disjoints et la séquence **quel qu'un**, état final, correspond à l'aboutissement de l'ensemble de ces opérations d'écriture où le seul écart qui subsiste est la déglutination du premier élément ;
- 6 - le dollar ou la suite de dollars (symboles \$, \$\$ ou \$\$\$) représente respectivement un caractère, une suite de caractères ou un mot d'un état antérieur dont les traces graphiques sont indéchiffrables partiellement ou en totalité. Ainsi (cf. l'exemple de la ligne 6), la forme **atteindre{\$\$\$}** renseigne que le verbe **atteindre** se substitue à une forme antérieure impossible à déchiffrer.
- Les lignes numérotées de 7 à 11 correspondent à différents types de marquages :
- 7 - la livre sterling (symbole £) figure une boucle en forme de e très ouvert dans le prolongement du jambage du dernier caractère d'un mot. Ainsi (cf. l'exemple de la ligne 7), la forme **coeur£**, que a interprétée la première fois comme **coeure** (en raison de la boucle en forme de e très ouvert) a fait croire à un écart graphique de l'auteur avec notation de -e en finale. Or, il n'en est rien. En réalité c'est une sorte de manie calligraphique, fonctionnant comme un ponctème de mise en valeur du mot ou de clôture de paragraphe. Ce signe est récurrent tout au long du volume et il est fréquemment situé à la terminaison de vedettes d'entrées d'articles ou à la fin de paragraphes ;
- 8 - le **trait souligné continu** mentionne tous les mots et expressions mis en relief par des effets graphiques dans le cours des articles (mise en valeur

matérialisée par une écriture droite épaisse s'opposant à une écriture fine penchée) ;

9 - le **trait souligné discontinu** signale tous les mots, expressions et additions situés dans les marges ;

10 - les crochets avec la mention BDR (symbole **[BDR]**) marquent tous les endroits où l'auteur se ménage un "blanc de réserve" destiné à recevoir une information ultérieurement. Cette information peut être le numéro d'une page de renvoi, une date dont les dizaines et les unités sont à préciser, un espace laissé libre pour finir la rédaction d'un article ou encore un espace laissé libre pour la rédaction d'un article dont seule la vedette d'entrée est écrite ;

11 - les crochets (symboles **[]**), encadrent tout signe, mot ou expression correspondant à une omission, qui est restitué. Il peut s'agir d'un signe de ponctuation omis, c'est souvent le cas de l'apostrophe. Cela peut être aussi utilisé pour signaler à l'attention le manque d'un mot, d'une expression ou d'un fragment de phrase dans un passage incompréhensible.

Ce système de marquage destiné à rendre compte de l'ensemble des éléments matériels n'est pour le moment que partiellement exploité, il fait l'objet d'une base de données en cours de constitution.

1.4. Modifications du manuscrit : comparaison de l'état initial à l'état final

À partir du marquage des écarts graphiques et des modifications du manuscrit, l'échantillon de la lettre A donne les résultats suivants :

Résultats globaux	Nombre total d'occurrences	15474	100,00 %
	État initial, nombre de formes sans écart	13957	90,19
	État final, nombre de formes sans écart	14247	92,07
Parties autographes	Nombre total d'occurrences	9846	100,00%
	Nombre total de formes sans écart	8872	90,10
Parties rapportées	Nombre total d'occurrences	5461	100,00%
	Nombre total de formes sans écart	5172	94,70
Parties modifiées	Pourcentage de réussite des substitutions	94,70 %	(sur 254)
	Pourcentage de réussite des additions	51,66 %	(sur 120)

• Ce tableau appelle quelques commentaires :

1 - tout d'abord, les résultats globaux montrent qu'entre l'état initial et l'état final l'auteur a réduit le nombre des écarts graphiques en passant d'un taux de réussite de 90,19 % à un taux de réussite de 92,07 % ;

2 - il ressort que les parties autographes présentent un taux de réussite de 90,10 % contre un taux de réussite de 94,70 % pour les parties rapportées. Cette différence montre simplement que l'auteur fait moins de fautes dans les opérations de reproduction de documents que dans le travail de rédaction des articles. Ici une étude complémentaire est nécessaire car il est fort probable que les écarts graphiques ne soient pas de même nature, selon qu'ils proviennent de l'une ou l'autre des parties.

- 3 - concernant les parties modifiées, la différence de réussite entre les substitutions et les additions s'explique par le fait que les additions correspondent dans de très nombreux cas à des actes d'hypercorrectisme, avec en particulier une tendance de l'auteur à mettre des -s finals de pluriel un peu partout (verbes à la 6^e personne : *les Normands ne **manqueronts** pas* (sous l'article *Arbre*, p. 68), adverbess, etc.).

Dans nos prochains travaux il est envisagé de confronter les résultats de ces modifications liées aux hésitations graphiques, aux formes dont les écarts existent dès l'état original. Une telle comparaison permettra certainement, à partir des points de fragilité de cette orthographe, de mettre en évidence ce qu'il est désormais possible d'appeler la *conscience graphique* de l'auteur.

2. Variation graphique

Avec Arsène Vincent nous sommes devant un paradoxe orthographique. En effet, comment peut-on expliquer son système graphique alors que dans le même article (*Année des Prussiens*), à quelques lignes d'intervalle, il réussit sur une opération d'accord syntaxique complexe et échoue sur un acte orthographique usuel ?

Ci-dessous, dans le premier exemple (p. 53 du *Dictionnaire*, en **gras** les écarts graphiques, en souligné les formes commentées) :

1	<i>Quelles frayeurs n' <u>ont-ils pas causées</u> à nos</i>
2	<i>mères & à leurs filles timides , ces militaires à l' oeil</i>
3	<i>menaçant , aux épaisses moustaches , & surtout ces grands</i>
4	<i>Cosaques au long sabre & à l' air farouche , <u>l'~ors~qu'</u> ils</i>
5	<i>entraient dans une ferme de village & que , d' un signe [...]</i>

- Réussite de l'accord du participe passé avec complément d'objet direct placé devant l'auxiliaire avoir (ligne 1, caractères soulignés) ;
- Deux écarts sur la segmentation de la conjonction *lorsque*, d'usage courant et soudée dès le XV^e-XVI^e siècle (ligne 4, caractères gras soulignés).

Ci-dessous, dans le second exemple (p. 56 du *Dictionnaire*) :

1	<i>lorsque <u>le père Bernier</u> leurs eut expliquer , <u>l{}</u>es</i>
2	<i>prussiens entrèrent tellement en colère d' avoir</i>
3	<i>manger de ce met , qu' ils prirent <u>le per<re> Bernier</u></i>
4	<i>au collet , l' amenèrent au poste (1) sur le champ</i>

- L'expression "le père Bernier" ne présente pas d'écart graphique (ligne 1, caractères soulignés) ;
- La même expression, 2 lignes plus bas, présente un écart graphique sur le mot *père* (ligne 3, caractères gras et soulignés). La graphie **per<re>** (état final), était phonétique à l'origine avec la forme **per**, et curieusement affublée après

coup d'un <re> (procédure d'addition), probablement par hypercorrectisme subconscient.

C'est cette accumulation de facteurs discordants concentrés dans le même article qui a amené à retenir celui-ci comme corpus de référence pour la présente communication. Cet article a pour titre *l'Année des Prussiens* (pp. 52-57) et présente les mêmes caractéristiques générales que l'ensemble de la lettre A (pourcentage de réussite état initial : 90,19 ; état final : 92,07). Le tableau des données, ci-dessous, confirme la représentativité de l'article choisi :

Nombre total d'occurrence	1066	
État initial, nombre de formes sans écart	956	pourcentage de réussite : 89,68
État final, nombre de formes sans écart	988	pourcentage de réussite : 92,68
Nombre de modifications	35	(7 ajouts et 28 substitutions)
Nombre d'écarts supprimés	32	

La suite du propos présente les écarts et les variantes graphiques les plus caractéristiques et les plus récurrents par ailleurs.

2.1. Secteur à dominante phonographique

Ce secteur concerne à la fois les écarts graphiques dans la notation de phonogrammes et la notation de variantes phoniques ou de dialectalismes. Mais toutes ces notations ne sont pas toujours systématiques, ce qui rend l'analyse plus délicate et certaines interprétations fragiles. Ces différents cas sont abordés ci-dessous :

- écarts et hésitations sur la notation de la nasale [ã]

La notation de cette voyelle nasale est problématique d'un bout à l'autre du *Dictionnaire*, mêmes écarts et hésitations au début, au milieu et à la fin de l'ouvrage, d'après les sondages effectués. Dans l'échantillon d'étude les variantes suivantes ont pu être observées (en italique gras les graphies de l'auteur) :

- ***demendaient*** pour *demandaient* (4 occurrences avec ***an***, contre 1 occurrence avec ***en***) ;
- ***anvirois*** pour *environs* (3 occurrences avec ***en***, contre 2 occurrences avec ***an***, dont 1 occurrence dans la partie Cah. Dol.) ;
- ***ma{e}nger*** pour *manger* (7 occurrences avec ***an***, contre 2 occurrences avec ***en***).

Sur l'ensemble des réalisations (19 formes), 14 graphies ne présentent pas d'écarts. L'usage normé est dominant. Il semble que l'explication de ces variations doit être recherchée dans l'analyse des contextes, en termes d'ensemble de tâches morphosyntaxiques à résoudre pour écrire selon la norme usuelle. Plusieurs hypothèses de recherche concernant ce type d'analyse sont abordées dans la conclusion.

- notation de *a* postérieur ?

L'accent circonflexe sur le *a* note-t-il le timbre postérieur de la voyelle ? La consultation de l'*ALIFO*⁵ permet de poser un doute. Thiron-Gardais qui est au centre du triangle défini par les points 32, 37 et 38 (cf. ci-dessous), donne respectivement, pour *mare*, deux réalisations antérieure de la voyelle contre une réalisation postérieure. Il semble que la voyelle est indifféremment réalisée selon l'une ou l'autre façon :

- ***mâre*** (*ALIFO*, T2, 368 "une mare" : 32 [mar], 37 [mar], 38 [mar], ***mares*** (dans la partie Rev. Agric.)

La forme ***mâre***, avec accent circonflexe, de l'article autographe (*Année des Prussiens*) s'oppose à la forme *mares*, sans accent, de la partie reproduite de la Rev. Agric., rien ne permet pour le moment d'affirmer si la forme autographe est une simple variante graphique ou si elle est aussi une variante phonique.

Par contre, pour *vase*, dont la forme locale est *bourbe*, la périphérie de l'aire de Thiron-Gardais atteste généralement la forme *vase* par *va* avec *a* postérieur (cf. ci-dessous), souvent concurremment avec la forme *bourbe* :

- ***vâse*** (*ALIFO*, T2, 373 "la vase de mare" : 32, 37, 38 [burb] ; 34 [va] et [burb], 17 [va], 25 [va] et [burb].

D'après la convergence de l'ensemble des réalisations, l'accent circonflexe de ***vâse*** pourrait être en rapport avec la notation du timbre postérieur de la voyelle.

- notation de yod ?

La forme ***millieu*** (article *Année des Prussiens*) atteste-t-elle un ancien *l* mouillé ou sa réduction à yod ? Il y a au total 3 occurrences du mot *milieu* sur totalité de la lettre A (cf. ci-dessous).

- ***millieu*** , p. 56 du *Dictionnaire* (TPP, p. 115, *miyeu* [Prosper Vallerange, 1861], article *Mitan*) ; André Gilbert, 1980-82, *miyeu*.

- ***milieu*** (2 occurrences, sous l'article *Arbres*).

Il faut remarquer que la graphie ***-illi-*** est très rare. Elle n'est attestée que 2 fois dans les entrées du *Dictionnaire* (sur 788 entrées) par la forme ***gadill<i>er*** "groseiller"⁶ et par la forme ***margouillia*** "flaque d'eau dans le chemin". La forme dialectale du *groseiller* a d'abord été écrite ***gadiller***, puis le ***i*** a été ajouté dans un second temps, peut-être par analogie avec la graphie *groseillier* (ce qui est représenté par ***i*** entre soufflets <*i*> dans la transcription), aboutissant ainsi à la graphie ***-illi-***. Dans le *Glossaire* ces mêmes entrées sont respectivement écrites par ***gadillier*** et ***margouilla*** (sans second ***i***). Ces variations graphiques, les hésitations de l'auteur et la rareté de la graphie ***-illi-*** ne permettent pas de considérer celle-ci comme une transcription de l'ancien *l* mouillé. Par contre le mot *milieu* étant souvent attesté avec yod sur l'aire dialectale (cf. *miyeu*, ci-dessus), cette graphie pourrait tout aussi bien être interprétée comme une

⁵ *ALIFO* = *Atlas linguistique et ethnographique de l'Ile-de-France et de l'Orléanais, Perche, Touraine*

⁶ *groseiller* (sans *i*) est la graphie rectifiée de *groseillier* (cf. *JO, Documents Administratifs*, n° 100 du 6 décembre 1990)

notation de yod, si l'on prend en considération l'insécurité graphique. À ces arguments il convient d'ajouter qu'Arsène Vincent s'applique à écrire selon la norme usuelle, avec une nette tendance à préférer la forme française à la forme dialectale.

- notation renforcée de *e* caduc passant à *e* fermé

L'accent aigu sur le second *e* d'*éperons* (cf. ci dessous) est une notation de l'oral.

- *épérons* pour *éperons*, mais aussi *ménusier* pour *menuisier* (sous l'article *Arbres*) ; Claire Fondet, 1980, *ménusier, tournévis, noisétier*.

Selon Claire Fondet 1995 « le patois (oral) résiste à la disparition de l'*e* par son passage à *é* ». La réalisation de cette voyelle neutre oscille entre zéro, [ə] et une voyelle à timbre [œ], [e] (Fondet 1980, 633).

2.2. Secteur à dominante morphographique

Ce secteur concerne généralement ce que l'on appelle d'ordinaire l'orthographe grammaticale. Si bon nombre de choix graphiques sont conditionnés par une analyse morphosyntaxique plus ou moins complexe (notation du genre, notation du nombre, morphologie du système verbal, segmentation des unités), d'autres choix graphiques sont plus ou moins inhérents aux aléas du système graphique lui-même (consonnes doublées en particulier). La conjugaison de ces facteurs produit chez le scripteur un sentiment d'insécurité orthographique qui est souvent à l'origine d'une forte tendance à l'hypercorrectisme. C'est un regard sur l'ensemble de ces facteurs qui est proposé ci-dessous.

- finales verbales en [E]

Le relevé exhaustif des finales verbales en [E] dans l'article de référence (*Année des Prussiens*), présente sur un total de 86 occurrences, 73 formes sans écart, contre 13 formes avec écarts. Le tableau, ci-dessous, représente de manière synthétique ces résultats :

	Imparfait		Infinitif		Participe passé	
Occurrences sans écart	43	-ait ou -aient	11	-er	19	-é, -ée et -és
Occurrences avec écart	2	-é	1	-é	9 1	-er -ait
Pourcentage de réussite	95,35		90,01		52,64	

Le tableau fait très nettement apparaître la corrélation qui existe entre le pourcentage de réussite et le degré de complexité de l'analyse morphosyntaxique. Par ailleurs, ce qui n'est pas complètement montré par le tableau, mais qui est bien mis en évidence par Marie-Rose Simoni-Aurembou (1989) c'est l'interchangeabilité de graphies qui traduisent une certaine agrammaticalité graphique de l'auteur. En effet, l'imparfait peut aussi présenter des finales graphiques en : *-er, -é, -és* (p. 16) ; l'infinitif des finales graphiques en : *-é, -ait, -ez* (p. 17) ; les participes passés des finales en : *-er, -ér, -ait* (p. 19).

Voir, ci-dessous des exemples, en contexte, de chacun des cas (formes en souligné gras) :

- imparfait :

*Beaudoux propriétaire du dit étang alla trouver leur chef qui était **loger** chez M. Taulé , & lui ayant **représenter** que s'y{il} il **laissé** [-ait] **pê{r}cher** l' étang **par{dans}** une **cha{el}leur** semblable*

. ici, le contexte montre bien qu'Arsène Vincent était vraisemblablement dépassé par la simultanéité des tâches à résoudre (gestion de participes passés, de l'imparfait, de l'infinitif et substitutions multiples dans la rédaction). D'autre part, l'écart graphique de la désinence de l'imparfait (dont le pourcentage de réussite est à 95,35 %), s'explique sans doute par le problème de la segmentation graphique (non résolu) placée juste devant le verbe.

- infinitif :

*Bernier , propriétaire à la **Coure-aux-Pauvres** , avait aussi deux **prussiens** à **hébergé** [-er]*

. écart graphique malgré la simplicité de la règle du verbe placé derrière la préposition.

- participe passé :

*ils voulurent savoir le nom de la viande qu' ils avaient **manger** [-ée] qui avait **était** [-é] **élevé{-e}** au **milieu** du champ de manoeuvre à la somme de 2000,500 francs*

. écart graphique probablement en rapport avec la complexité de la règle d'accord du complément d'objet direct placé avant l'auxiliaire avoir. Le second cas, participe passé du verbe *être*, semble relativement récurrent par une sorte d'analogie graphique des formes verbales (cf. Simoni-Aurembou, 1989 : 19).

Tous ces exemples semblent indiquer que l'analyse doit prendre en compte à la fois la complexité des règles morphosyntaxiques et les problèmes de surcharge cognitive qui s'accumulent en certains endroits.

• segmentation particulière

L'insécurité orthographique d'Arsène Vincent se manifeste aussi par la segmentation variable de la chaîne morphosyntaxique des unités graphiques. Dans le tableau (*variantes dans même source*) présenté supra, pour la conjonction *lorsque* il ne propose pas moins de trois déglutinations différentes de cette unité : *l' orsque* et *lors que*, en deux éléments graphiques ; et *l' ors qu'* en trois éléments graphiques. (cf. ci-dessous) :

1	<i>Cosaques au long sabre & à l' air farouche , l'~•ors~•qu' ils</i>
2	<i>entraient dans une ferme de village & que , d' un signe</i>

Avec cette conjonction sont posées la plupart des difficultés de la reconnaissance et de la délimitation graphique des unités, à partir de l'oral. Cette variation de la segmentation concerne essentiellement les mots-outils les plus

fréquents⁷ (articles, pronoms, conjonctions, adverbes, locutions, et formes verbales de haute fréquence d'usage).

Dans l'exemple ci-dessous, une première agglutination (représentée par le dièse #) de la locution *à peine* est suivie par une seconde agglutination, à la ligne suivante, de la conjonction *que* avec le pronom *ils* (en **gras** les écarts graphiques, en souligné les formes commentées) :

1	<i>disposition . Deux heures après , <u>a#peine</u> ci l'] eau</i>
2	<i>é{l-}tait écoulée <u>qu#ils</u> sautèrent dans la vâse en y</i>

La séquence graphique *qu#<'>e<n>* met en évidence les modifications successives visibles sur le manuscrit (cf. ci-dessous, ligne 2). Dans un premier jet Arsène Vincent écrit *que*, probablement entraîné par la construction syntaxique. Dans un second temps l'insertion de la préposition *en*, devant la date, se traduit par l'insertion d'un *n* à la suite du *e* de *que* qui devient ainsi une unité graphique sous la forme *quen*. Dans un troisième temps enfin, il insère une apostrophe entre *qu* et *en*, pour séparer l'unité de la séquence en deux éléments graphiques : *qu'en*.

1	<i>partie du poisson perdu ; Cette bonde ne fut rétablie</i>
2	<i><u>qu#<'>e<n></u> 1861 .</i>

L'agglutination graphique d'*entre autres*, en une forme ***entrautres*** (voir ci-dessous), est peut être influencée par les modèles des substantifs du type *entr'acte* ou des verbes pronominaux du type *s'entr'aider* en usage dans le *Dictionnaire de l'Académie française* de 1835 ? D'après cette hypothèse, et l'ajout à posteriori de l'apostrophe comme montré à l'exemple précédent (***qu'en***), il est vraisemblable de déduire qu'Arsène Vincent a oublié de mettre l'apostrophe, signe qui est souvent posé après l'écriture du mot. Cette interprétation paraît tout à fait fondée si elle est rapprochée de l'autre occurrence ***entre'autre*** (sans -s final, p. 59), située 2 pages plus loin, où le mot est écrit en une seule séquence graphique avec l'apostrophe ajoutée dans un second temps.

1	<i>en emportant beaucoup de buttin aux habitants de</i>
2	<i>Thiron , <u>entr#autres</u> , une paire d' épérons en argent ,</i>
3	<i>une montre en or & une paire de boucle d' oreil , avec</i>

- tendance à l'hypercorrectisme

L'insécurité orthographique d'Arsène Vincent se traduit encore par une tendance à l'hypercorrectisme :

- par accentuation :

- . ***fût***, avec accent circonflexe (3 occurrences), 3^e personne du passé simple pour *fut*

⁷ Cf. N. Catach, F. Jejeic et équipe HESO, *Les listes orthographiques de base*, Annexe 1, pp. 41-46.

cette manie de l'accentuation, en dehors des formes sans accent qui en reçoivent un comme par exemple *loûtre*, montre la préférence accordée à l'usage de l'accent circonflexe, qui avec des formes telles que *dîre*, *noël*, *zèle*, *vîte*, acquiert une valeur symbolique synonyme d'érudition.

- par doublement de consonnes :

. *buttin* pour *butin*, avec une occurrence réduite à *t* simple dans *but{t}in*

. *lannières* pour *lanières*

. *parrages* pour *parages*

. *trippes* pour *tripes*

cette tradition trouve sans doute son origine dans le système orthographique du français dont les nombreuses lettres superflues et non fonctionnelles induisent ce comportement irrationnel du scripteur.

- par ajout de marques de nombre :

. *ens* pour *en*, pronom, apparaît dans le contexte : « *quoiqu' ils s' ens furent bien régaler* ». Ici, le contexte pluriel de l'énoncé semble avoir incité l'auteur à ajouter un -s final, marque de pluriel, à *en*..

. *leurs* pour *leur*, adjectif ou pronom, survient dans des contextes du pronom personnel complément d'objet indirect ou dans des contextes se rapportant à plusieurs possesseurs. Arsène Vincent maîtrise mal ces règles, comme le montrent les tournures analogues des exemples suivants, à seulement une page d'intervalle : « *se qui était à leurs convenance* », p. 54 / « *ce qui pouvait être à leur convenance* », p. 53. Le pronom personnel invariable, qui renvoie à des situations plus abstraites, présente un plus faible de taux de réussite (cf. ci-dessous) :

- adjectif : 8 occurrences, 2 écarts, pourcentage de réussite 84 %

- pronom : 4 occurrences, 2 écarts, pourcentage de réussite de 50 %

• notation du nombre

- dans composés au pluriel introduit par *d(e)* ou dans des locutions :

. *boucle d'oreil* pour *boucles d'oreille*, apparaît dans le contexte : « *une paire de boucle d' oreil* ». Comment expliquer l'absence de marques graphiques du pluriel ? Une première interprétation peut s'expliquer par le contexte antérieur au singulier avec préposition invariable, élidée devant substantif à initiale vocalique. Une seconde interprétation sur l'absence de marques peut trouver son origine dans une transcription de l'oral. En effet, dans un autre contexte sous l'article *Arbres* : « *le pommier est d'une grande utilité pour faire les oreils & les oreillons de charrue* », le [z] de liaison qui se fait entendre aux initiales vocaliques semble avoir entraîné la notation par -s du pluriel, cette dernière interprétation paraît plausible étant donné la récurrence dans l'oral du syntagme [zɔrej].

. *toute nation* pour *toutes nations* apparaît dans le contexte : « *Des soldats de toute nation* ». Cet écart est fréquent dans les expressions globales de la

totalité, comme ici, ou dans des expressions distributives de la totalité comme par exemple « *chaque **pompier**s* », avec -s final à *pompier*, sous l'article *Aguilan*.

- notation du genre
 - même écart d'accord dans le texte (ligne 1 à 4) et dans la note de bas de page (lignes 5 et 6), cf. ci-dessous, en **gras** les écarts graphiques, en souligné les formes commentées :

1	<i>Dans ce bataillon se trouvait un prêtre</i>
2	<i>hollandais , il fut dire la messe sur <u>une au{ta-}tel</u> (2)</i>
3	<i>en terre et planche , qui avait <u>était élevé{-e}</u> au</i>
4	<i>milieu du champ de manoeuvre qui était le</i>
5	<i>(2) Cet autel avait été <u>élevé{-e}</u> au frais de la commune de</i>
6	<i>Thiron , & par les mains des enfants Doirat , maçons à Thiron .</i>

Cette exemple montre un contexte de modifications où l'auteur devait gérer à la fois une substitution lexicale avec changement de genre (substitution de *table* par *autel*), un appel de note de bas de page avec un texte issu de la même substitution lexicale, avec changement de genre, et un participe passé en rapport avec cette substitution. Il semble ici que l'enchaînement des tâches lié au changement de genre du lexème soit à l'origine d'un problème cognitif qui peut expliquer l'oubli du changement de genre de l'article indéfini, *une* > *un*, dans l'expression « ***une autel*** ».

- écart d'accord lié à la distance entre les unités :

L'adjectif ***perdu*** a été spontanément accordé avec *poisson*, substantif le plus proche, interposé entre le substantif féminin et l'adjectif avec lequel doit se faire l'accord.

1	<i>de son propriétaire , la bonde en fût coupée & une</i>
2	<i>partie du poisson perdu ; Cette bonde ne fut rétablie</i>
3	<i>qu#<'>e<n> 1861 .</i>

- francisation (hors article *Année des Prussiens*)

Bien que n'ayant aucune attestation de francisation de formes dialectales dans le corpus de la présente étude, il est important de signaler l'existence de ce phénomène qui dénote chez Arsène Vincent son aspiration à montrer qu'il connaît et écrit la variété haute de la langue. D'ailleurs, sous l'article dont l'entrée est ***Féliau*** (forme dialectale), il ne manque pas de faire remarquer qu'il s'agit du « *nom corrompu de **fléau**, instrument pour battre le blé* » (*Dictionnaire*, p. 337).

Il semble bien que ce soit cette disposition d'esprit qui est à l'origine des deux graphies ***ferseau*** et ***avaloire***, cf. ci dessous :

- **ferseau**, est attesté dans le *Dictionnaire*, p. 72, par l'expression « **ferseau de charrues** », sous l'article *Arbres*, dont la forme dialectale correspondante est [fersjo] (*ALIFO*, T1, 73, "têtard (de la charrue)", partie de l'avant-train). La forme [fersjo] est attestée aux points 32, 37 et 38 de l'*ALIFO* dont Thiron-Gardais est le centre. La graphie **ferseau** apparaît donc comme une forme francisée où le suffixe dialectal [-jo], type **biau** pour *beau*, passe à [o] par l'intermédiaire du graphème *-eau*.

- **avaloire**, est une entrée d'article du *Glossaire*, p. 2, pour **avalouère** (*TPP*, p. 38, "pièce de harnais, gourmand").

La vedette d'article **avaloire**, est comme pour l'exemple précédent, une francisation de la forme dialectale **avalouère** qui aurait dû être la forme d'entrée attendue du *Glossaire*.

Conclusion

L'état actuel des recherches, n'a pas encore permis de mettre en évidence les systématiques sous-jacentes au système graphique d'Arsène Vincent. Les facteurs de la variation graphique, en raison de l'insécurité orthographique de l'auteur, sont extrêmement instables et difficiles à isoler de manière certaine. À travers les hypothèses et les interprétations présentées ici, un certain nombre de pistes de recherche sont ouvertes. La plus fructueuse paraît être celle des aspects cognitifs de l'écrit, à la fois dans les représentations de la langue et de ses variétés, comme dans la réalisation matérielle de la mise en écrit de celles-ci. La constitution en cours de la base de données sur l'ensemble des modifications du manuscrit, comme ensemble organisé des traces de tâtonnements du scripteur, paraît être un moyen privilégié pour avancer dans ce travail et dans la connaissance des processus d'écriture à partir des témoignages historiques de scribes peu assurés des campagnes françaises au milieu du XIX^e siècle.

Fabrice Jejcic

CNRS, Paris, "Histoire des systèmes graphiques du français et de ses variétés"

Annexe 1

Données sur le Dictionnaire et le Glossaire

- **Dictionnaire**

Le **Dictionnaire** est inachevé, sa nomenclature s'étend de la lettre **A** à la lettre **P**. Le **volume** est composé de 55 cahiers, 12 pages pliées en deux, le papier est épais et de bonne qualité.

La **pagination** comporte 718 pages au total, dont 694 sont remplies d'une écriture manuscrite fine et très lisible.

La **datation** est basée sur des sondages effectués au début, au milieu et à la fin du volume. Ce qui permet de penser que le volume a été composé vers la fin des années 1870, peu de temps avant la mort d'Arsène Vincent survenue en 1881.

Le **contenu** du Dictionnaire s'articule autour de rubriques traitant de l'histoire, des coutumes et du langage du canton de Thiron-Gardais. Ces rubriques sont classées alphabétiquement.

La **structure** du Dictionnaire s'appuie sur de courtes **gloses** autographes de vocables patois (généralement de quelques lignes) et sur des articles de type **encyclopédique** avec parties autographes et parties rapportées provenant d'autres sources toujours citées (la longueur de ces articles varie de la demi-page à plusieurs dizaines de pages).

- **Glossaire**

La nomenclature du **Glossaire** s'étend de la lettre **A** à la lettre **V**. La **pagination** comporte 16 pages manuscrites. La datation du Glossaire le situe en 1864, d'après la date trouvée dans l'article *épine-blanche*.

- **Comparaison des nomenclatures :**

- nombre total d'articles du **Dictionnaire** : **788** (environ *950 par extrapolation)
- nombre total d'articles du **Glossaire** : **429** (356 jusqu'à la lettre P, différence de 73)
- nombre total de vedettes **communes** : **211**
- nombre total de vedettes **différentes** : **933**

- **Données rapportées au corpus retenu de la lettre A :**

- nombre total d'articles du **Dictionnaire** : **71** (dont **39** mots du cru)
- nombre total d'articles du **Glossaire** : **35** (dont **28** mots du cru)

Annexe 2

**Bref aperçu des contrastes et de la variation graphique :
articles du *Dictionnaire* / articles du *Glossaire*,
en contraste avec les entrées du *Trésor du parler percheron* [TPP]**

<p>Avril . - <i>l'orsqu'il tombe une petite pluie d{ }ouc{s}e, soit en mars soit en avril, on dit : il avril, tout va sortir de terre . _ voyez poisson d'avril, page (Dictionnaire, p. 82)</i></p> <p>Avrilé . - <i>Quand il tombe une petite pluie fine au mois d'avril, on dit : il avrile (Glossaire, p. 2)</i></p> <p>Avriller (TPP, p. 38)</p>
<p>Achant . - <i>actif à l'ouvrage, (Dictionnaire, p. 4)</i></p> <p>Hachant . - <i>personne active, leste & aimant le travail{le}. (Dictionnaire, p. 424)</i></p> <p>hachant . - <i>Etre <u>achant</u> à l'ouvrage, qui à du courage, vif. - (Glossaire, rien sous A, p. 9)</i></p> <p>Achant (TPP, Vi., p. 38)</p>
<p>Aise & aisiau - <i>On donne ce nom aux barrières de jard{a}in . Ce nom vient de ce qu' autrefois il y avait à chaque portes d' une habitation , une seconde porte en bois don le haut était denté comme une scie , & que l' on appelait <u>ais</u> . (Dictionnaire, p. 48)</i></p> <p>Aise, ou Aisiau . - <i>On donne se nom au contre-porte en paille que lon fait à l['] hiver pour se garentir du froid. (Glossaire, p. 1)</i></p> <p>Haise, "clôture légère" (TPP, dont Vi., p. 100)</p> <p>Haisieau, "contre-porte en paille" (TPP, dont Vi., p. 100)</p>
<p>Féliau£ . - <i>nom corrompu de <u>fléau</u>, instrument pour battre le blé [...](Dictionnaire, p. 337)</i></p> <p>Félliot . - <i><u>Fléau</u> pour batre le grain [...] (Glossaire, p. 7)</i></p> <p>Féliau (TPP, dont Vi., p. 86)</p>

Bibliographie succincte

- DUD'HUIT, Albert, MORIN, Alain, SIMONI-AUREMBOU, Marie-Rose, 1996, *Trésor du parler percheron*, Cahiers Percherons, Numéro hors série, Fédération des amis du Perche, réédition, 318 p. [1^{ère} éd. 1979].
- FONDET, Claire, 1980, *Dialectologie de l'Essonne et de ses environs immédiats*, Lille-Paris, Université de Lille III-Librairie Honoré Champion, 2 tomes, 746 p.
- FONDET, Claire, 1995, "La tradition orale chez André Gilbert", pp. XIX-XXII, dans *L'héritage perdu ou l'amour de la terre*, par André Gilbert, Dijon, ABDO, 202 p.
- PONTOIRE, Jean, FONDET, Claire, éd., 1999, *Glossaire des parlers d'Eure-et-Loir, Beauce et Perche, d'après l'enquête de 1868*, Chartres, Société Archéologique d'Eure-et-Loir (SAEL), 318 p.
- SIMONI-AUREMBOU, Marie-Rose, 1973, *Atlas linguistique et ethnographique de l'Ile-de-France et de l'Orléanais, Perche, Touraine*, tome 1, cartes Paris, CNRS, cartes 1 à 318 p.
- SIMONI-AUREMBOU, Marie-Rose, 1978, *Atlas linguistique et ethnographique de l'Ile-de-France et de l'Orléanais, Perche, Touraine*, tome 2, Paris, Éditions du CNRS, cartes 319 à 387.
- SIMONI-AUREMBOU, Marie-Rose, 1982, *L'alphabet du quotidien, petite chronique du canton de Thiron au XIX^e siècle*, Cahiers percherons 69-70, Mortagne-au-Perche (28), Association des Amis du Perche, 60 p.
- SIMONI-AUREMBOU, Marie-Rose, 1989, "L'orthographe d'un charron d'Eure-et-Loir au XIX^e siècle", pp. 13-26, dans *La variation de la langue en France du XVI^e au XIX^e siècle*, Paris, Éditions du CNRS, 136 p.